Genetic Markers: DNA sequences with known locations on chromosomes.

- Points of variation that help:
  - Identify individuals or species
  - Associate inherited disease with nearby unidentified genes
- The availability of genetic markers is directly related to the presence of **polymorphisms**.

Polymorphism: Occurrence of two or more DNA sequences (alleles) at a locus among individuals, groups or populations.

Examples: - Single nucleotide polymorphisms (SNPs)
- Minisatellites

https://www.nature.com/subjects/genetic-markers

Single-nucleotide Polymorphism (SNPs)

- Result from a single base substitution
- Much more frequent in human genome than micro satellites
- Occur in introns, exons, promoters and intergenic regions

Fingerprinting (Microsatellite)

Most common genetic markers used for linkage analysis are **Microsatellite** markers.

- Concentrate on parts of the genome that are highly varied.
  - Short Tandem Repeats (STRs)
  - Places in genome with some **motif**

Sequence motif: recurring patterns in DNA that are ~~presumed to have a biological function~~.

- Repeated many times (2-6 bp repeat units)
- Highly variable: exceptionally high mutation rates(sometimes as high as 1 in 10)
- Why so high? Not sure; weak evidence
- Amplified by PCR
  - Automated screening **wing** fluorescent primers and laser detection
  - Different alleles differ in number of repeats, sometimes amplified sections will be longer, sometimes shorter
    - To find out how long a particular repeat N, run it through column, which ***retards*** DNA fragment according to length
    - Long out to come out column = longer fragment
    - Results presented like electrophoresis gel
    - Each column represents different individual
    - Different colours in columns represent results of different PCR reactions
      1. single yellow band - individual is homozygous at that particular locus
      2. two yellow bands - individual is heterozygous

Background: - Mauritius Kestrel nearly died out
- Intensive breeding programme to try and save them = 400 new breeding pairs
- Had genetic samples from before crash in numbers -> right hand side results
  - Same bands on right side and left side lost
  - Individuals on left hand side more likely to be homozygous

- At each locus, one allele very common so quite often individual will have two copies of that allele

Mini vs Microsatellite

-Mini satellites have longer repeat units (1 **ub** or more)

Hardy-Weinberg Principle Assumptions

- Organisms are diploid
- Generations are non-overlapping in diploid sexual organism
- Population must be large
- Non immigration or emigration
- No mutation in the gene of interest
- No natural selection (individuals reproduce at equal rates)
- Mating is random
- Negligible mutation, migration and selection

- If gametes combine at random, HW equilibrium as observed in 1 generation

How HW can be used in real applications - Chi-squared test

- Genotype —> allele frequency count
- Use HW to calculate expected frequencies of each variant
- To find out if results due to sampling variation - do a test
    - Calculate the difference between observed and expected values
    - Standardise by dividing by expected value
- Genetic drift has potential to produce larger differences than sampling variations
- If large differences are found in snail populations, they can be attributed to sampling variation, might be due to genetic drift and/or selection

Hardy-Weinberg Principle

- ALL ABOUT FREQUENCIES

"frequencies of alleles & genotypes in a population will remain constant over time in absence of other evolutionary influences"

① Frequency of Dominant allele	p
Frequency of recessive allele	q

① The sum of all possible outcomes must equal 1. $p + q = 1$

NOTE:
[frequencies of ~~recessive~~ + dominant must equal one because have to have 1 or the other if 2 alleles involved]
* since both equal to 1, inequality can be formed.

Genotype Frequency

is:

Allele #1 $p + q = 1$ AND Allele #2 $p + q = 1$

means multiply

$(p+q)(p+q) = 1$

KEY

$p^2 + 2pq + q^2 = 1$

$p^2 + 2pq + q^2 = 1$

- p^2 — Homozygous Dominant — R R
- $2pq$ — Heterozygous — R r
- q^2 — Homozygous recessive — r r

Problem with Sampling Variation

- Sometimes large deviation from expected results, sometimes small deviation
- Statisticians can tell how often sampling variation will be large or huge
- If the null hypothesis were true (if it is just sampling variation producing deviation from results), how often will we see a result this extreme?
    - That is what P Value tells us
- P value doesn't tell us how likely it is that null hypothesis is true, also doesn't tell us how likely result is due to selection

Wahlund Effect

- Is the increased frequency of homozygotes in subdivided populations
- If two or more subpopulations have different allele frequencies then the overall heterozygosity is reduced, even if subpopulations themselves are in HW equilibrium
- Underlying causes: could be geographical barriers to gene flow followed by genetic drift in subpopulations

Example:

1. Population 1: high A frequency, predominantly AA
2. Population 2: high B frequency, predominantly BB

Heterozygosity reduced

Wahlund Effect: Consequences

- If there are more homozygotes than expected from HW principal, it may lead top suspect that the selection was favoring homozygotes, when an actual fact both subpopulations are in HW equilibrium and the deviation is actually due to unintended pooling of separate populations.
    - It is therefore important to know the structure of a population when applying HW principle.
- When a number of previously subdivided populations merge together, the frequency of homozygotes will decrease.
    - In humans this leads to decrease in incidence of rare recessive genetic diseases when previously isolated population comes together with larger population →decrease only expressed in homozygous and when two populations interbreed the frequency of homozygotes goes down.

Background on Mauritius Kestrel Breeding Programme

https://www.newscientist.com/article/mg23130891-100-the-mauritius-kestrel-is-living-proof-lets-change-conservation/

- Estimated 2 breeding pairs left in wild, some in captive
- Captives died due to DDT-ridden mice
- Years of conservation work by carl jones – 300+ birds now on island

Text

http://onlinelibrary.wiley.com/doi/10.1046/j.1365-2656.2001.00502.x/full

- Results show 2 breeding pairs accurate – highly severe bottleneck

Genetic Variation in Mauritius Kestrels

https://s3.amazonaws.com/academia.edu.documents/46409199/Nichols_RA_Bruford_MW_Groombridge_JJ._Su20160611-12337-190d1x4.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1508855780&Signature=eT7EuFXp6TKo5FJtohQxfwtVwbI%3D&response-content-disposition=inline%3B%20filename%3DSustaining_genetic_variation_in_a_small.pdf (copy&paste)

- High genetic diversity in museum specimens (before bottleneck)
- Restored vs. ancestral – reduced observed heterozygosity (0.099 vs. 0.231)
- Current diversity higher than expected
- Current populations have poor dispersal – no detected mixing
- This could have helped maintain current diversity (compared to undivided pop.)

Variation in above/below ground mosquitos

https://www.nature.com/hdy/journal/v82/n1/full/6884120a.html

- *molestus* form of *C. pipiens* live & breed in London underground
- morphologically the same as *C. pipiens* but different ecology and behaviour (https://www.nature.com/hdy/journal/v82/n1/fig_tab/6884120t1.html#figure-title)
- *molestus* form has low heterozygosity
- all *molestus* breeding crosses produced F1 which went on to produce F2 generations
- gene flow between *pipiens* & *molestus* absent or reduced, even when in contact
- significant genetic differentiation between lines – tunnel system

Hardy-Weinberg Principle

https://www.nature.com/scitable/knowledge/library/the-hardy-weinberg-principle-13235724
https://en.wikipedia.org/wiki/Hardy%E2%80%93Weinberg_principle

Chi-squared Test

http://www.statisticshowto.com/probability-and-statistics/chi-square/
https://en.wikipedia.org/wiki/Chi-squared_test